On The Ordering of the Spectral Radius Product $r(\mathbf{A})r(\mathbf{AD})$ Versus $r(\mathbf{A}^2\mathbf{D})$ and Related Applications

Lee Altenberg

January 22, 2013

Abstract

For a nonnegative matrix $\mathbf{A}$ and real diagonal matrix $\mathbf{D}$, two known inequalities on the spectral radius, $r(\mathbf{A}^2\mathbf{D}^2) \geq r(\mathbf{AD})^2$ and $r(\mathbf{A})\, r(\mathbf{AD}^2) \geq r(\mathbf{AD})^2$, leave open the question of what determines the order of $r(\mathbf{A}^2\mathbf{D}^2)$ with respect to $r(\mathbf{A})\, r(\mathbf{AD}^2)$. Here, sufficient conditions are found on $\mathbf{A}$ that determine orders in either direction. For a symmetrizable nonnegative matrix $\mathbf{A}$ with all positive eigenvalues and nonnegative $\mathbf{D}$, $r(\mathbf{A}^2\mathbf{D}) \leq r(\mathbf{A})\, r(\mathbf{AD})$. The reverse holds if all of the eigenvalues of $\mathbf{A}$ are negative besides the Perron root. This is a particular case of the more general result that $r(\mathbf{A}[(1-m)\, r(\mathbf{B})\, \mathbf{I} + m\mathbf{B}]\mathbf{D})$ is monotonic in m when all non-Perron eigenvalues have the same sign — decreasing for positive signs and increasing for negative signs, for symmetrizable nonnegative $\mathbf{A}$ and $\mathbf{B}$ that commute. Commuting matrices include the Kronecker products $\mathbf{A}, \mathbf{B} \in \{\otimes_{i=1}^{L} \mathbf{M}_i^{t_i}\}$, $t_i \in \{0, 1, 2, \ldots\}$, which comprise a class of application for these results. This machinery allows analysis of the sign of $\partial/\partial m_j\, r(\{\otimes_{i=1}^{L}[(1-m_i)\, r(\mathbf{A}_i)\mathbf{I}_i + m_i\mathbf{A}_i]\}\mathbf{D})$. The eigenvalue sign conditions also provide lower or upper bounds to the harmonic mean of the expected sojourn times of Markov chains. These inequalities appear in the asymptotic growth rates of viral quasispecies, models for the evolution of dispersal in random environments, and the evolution of site-specific mutation rates over the entire genome.

Keywords: sojourn time, reversible Markov chain, positive definite, conditionally negative definite

MSC Subject Classification:

15A42 Inequalities involving eigenvalues and eigenvectors,
15B48 Positive matrices and their generalizations,
15A18 Eigenvalues, singular values, and eigenvectors.

1 Introduction

In a recent paper, Cohen [9] compares two inequalities on the spectral radii, r, of products involving nonnegative square matrix $\mathbf{A}$ and positive diagonal matrix $\mathbf{D}$:

$$r(\mathbf{A}^2\mathbf{D}^2) \geq r(\mathbf{AD})^2, \tag{1}$$

$$r(\mathbf{A})\, r(\mathbf{AD}^2) \geq r(\mathbf{AD})^2. \tag{2}$$

Inequality (1) is obtained in [10], while inequality (2) is obtained in [9]. The relationship between two the left-hand side expressions in the inequalities is not determined. Cohen notes that positive matrices $\mathbf{A}$ and real diagonal matrices $\mathbf{D}$ can be chosen to give either

$$r(\mathbf{A}^2\mathbf{D}^2) > r(\mathbf{A})\, r(\mathbf{AD}^2) \quad \text{or} \quad r(\mathbf{A}^2\mathbf{D}^2) < r(\mathbf{A})\, r(\mathbf{AD}^2), \tag{3}$$

and asks whether conditions may be found that guarantee a direction to the inequality.

This seemingly narrow question is intimately related to a very broad class of problems that arise repeatedly in ecological and evolutionary dynamics [3, 5, 4]. Specifically, it relates to the open question of what conditions on $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$ determine the sign of $\mathrm{d}r([(1-m)\mathbf{A} + m\mathbf{B}]\mathbf{D})/\mathrm{d}m$.

Two results from the population biology literature provide conditions that determine the direction of inequality in (3) when $\mathbf{A}$ is a stochastic matrix. Both results rely on the constraint that the stochastic matrix be *symmetrizable*:

Definition 1 (Symmetrizable Matrix)**.** *A square matrix* $\mathbf{A}$ *is called* symmetrizable *to a symmetric matrix* $\mathbf{S}$ *if it can be represented as a product* $\mathbf{A} = \mathbf{D}_{\mathsf{L}}\mathbf{S}\mathbf{D}_{\mathsf{R}}$*, where* $\mathbf{D}_{\mathsf{L}}$ *and* $\mathbf{D}_{\mathsf{R}}$ *are positive diagonal matrices.*

Conditions for the inequality $r(\mathbf{A}^2\mathbf{D}) \leq r(\mathbf{A})\, r(\mathbf{AD})$ are found in the following theorem of the late Sam Karlin.

Theorem 2 ([15, Theorem 5.1, pp. 114–116, 197–198])**.** *Consider a family* $\mathcal{F}$ *of stochastic matrices that commute and are simultaneously symmetrizable to positive definite matrices, i.e.:*

$$\mathcal{F} := \{\mathbf{M}_i = \mathbf{D}_{\mathsf{L}}\mathbf{S}_i\mathbf{D}_{\mathsf{R}} \colon \mathbf{M}_h\mathbf{M}_k = \mathbf{M}_k\mathbf{M}_h\}, \tag{4}$$

where $\mathbf{D}_{\mathsf{L}}$ *and* $\mathbf{D}_{\mathsf{R}}$ *are positive diagonal matrices, and each* $\mathbf{S}_h$ *is a positive definite symmetric nonnegative matrix. Let* $\mathbf{D}$ *be a positive diagonal matrix. Then for each* $\mathbf{M}_h, \mathbf{M}_k \in \mathcal{F}$*:* $r(\mathbf{M}_h\mathbf{M}_k\mathbf{D}) \leq r(\mathbf{M}_k\mathbf{D})$.

The theorem applies to (3) by constraining $\mathbf{D}^2$ to be a positive diagonal matrix and substituting $\mathbf{M}_h = \mathbf{M}_k = \mathbf{A}$, and $r(\mathbf{A}) = 1$, which yields $r(\mathbf{A}^2\mathbf{D}^2) \leq r(\mathbf{A})r(\mathbf{AD}^2)$.

Theorem 2 is extended in [3] to conditions that make the spectral radius monotonic over a homotopy from $\mathbf{M}_k$ to $\mathbf{M}_h\mathbf{M}_k$. This monotonicity, either increasing or decreasing, establishes inequalities in each direction between $r(\mathbf{A}^2\mathbf{D})$ and $r(\mathbf{A})\, r(\mathbf{AD})$.

Theorem 3 (From [3, Theorem 33]). *Let $\mathbf{P}$ and $\mathbf{Q}$ be transition matrices of reversible ergodic Markov chains that commute with each other. Let $\mathbf{D} \neq c\,\mathbf{I}$ be a positive diagonal matrix, and define*

$$\mathbf{M}(m) := \mathbf{P}[(1-m)\mathbf{I} + m\mathbf{Q}], \qquad m \in [0,1]. \tag{5}$$

If all eigenvalues of $\mathbf{P}$ are positive, then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m < 0$. If all eigenvalues of $\mathbf{P}$ other than $\lambda_1(\mathbf{P}) = 1$ are negative, then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m > 0$.

The condition in Theorem 2 that the matrices be symmetrizable is shown in [2, Lemma 2] to be equivalent to their being the transition matrices of reversible Markov chains. Theorem 3 yields Theorem 2 by letting $\mathbf{M}_k = \mathbf{P}$ and $\mathbf{M}_h = \mathbf{Q}$. Then $\mathbf{M}(0) = \mathbf{M}_k$ and $\mathbf{M}(1) = \mathbf{M}_h\mathbf{M}_k = \mathbf{M}_k\mathbf{M}_h$. The hypothesis that $\mathbf{M}_k$ is symmetrizable to a positive definite matrix means that $\mathbf{M}_k$ has all positive eigenvalues, so $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m \leq 0$ for any positive diagonal $\mathbf{D}$, and thus $r(\mathbf{M}_h\mathbf{M}_k\mathbf{D}) = r(\mathbf{M}_k\mathbf{M}_h\mathbf{D}) \leq r(\mathbf{M}_k\mathbf{D})$. Note that the eigenvalues of $\mathbf{M}_h$ are generally irrelevant to this inequality.

For the inequality in the reverse direction, $r(\mathbf{A}^2\mathbf{D}) \geq r(\mathbf{A})\, r(\mathbf{AD})$, let all the eigenvalues of $\mathbf{A}$ other than $r(\mathbf{A}) = \lambda_1(\mathbf{A}) = 1$ be negative and substitute $\mathbf{A} = \mathbf{P} = \mathbf{Q}$, so $\mathbf{M}(0) = \mathbf{A}$ and $\mathbf{M}(1) = \mathbf{A}^2$. The result $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m \geq 0$ yields $r(\mathbf{A}^2\mathbf{D}) \geq r(\mathbf{A})r(\mathbf{AD})$ for such a stochastic symmetrizable matrix $\mathbf{A}$.

In the present paper, Theorem 3 is generalized to all symmetrizable irreducible nonnegative matrices.

2 Results

The following notational conventions are used. The elements of a matrix $\mathbf{A}$ are $[\mathbf{A}]_{ij} \equiv A_{ij}$, the columns are $[\mathbf{A}]_i$, and the rows are $[\mathbf{A}]^i$. A diagonal matrix with elements of a vector $\mathbf{x}$ along the diagonal is $\mathbf{D}_\mathbf{x} = \mathbf{diag}\big[\mathbf{x}\big]$. When $\mathbf{D} \neq c\,\mathbf{I}$ for any $c \in \mathbb{R}$, $\mathbf{D}$ is called *nonscalar*. The vector with 1 in position i and zeros elsewhere is $\mathbf{e}_i$.

We review the properties of irreducible nonnegative $n \times n$ matrices. When $\mathbf{A}$ is irreducible then for each (i,j) there is some $t \in \mathbb{N}$ such that $[\mathbf{A}^t]_{ij} > 0$. The eigenvalues of $\mathbf{A}$ are represented as $\lambda_i(\mathbf{A})$, $i = 1, \ldots, n$, and the spectral radius by $r(\mathbf{A}) := \max_{i=1,\ldots,n} |\lambda_i|$. We recall from Perron-Frobenius theory that $r(\mathbf{A})$ is a simple eigenvalue of $\mathbf{A}$, called the *Perron root*, designated here as $r_A \equiv r(\mathbf{A}) = \lambda_1(\mathbf{A})$. The *non-Perron* eigenvalues are $\lambda_{Ai} \equiv \lambda_i(\mathbf{A})$, $i = 2, \ldots, n$. Let $\mathbf{v}(\mathbf{A})$ and $\mathbf{u}(\mathbf{A})^\top$ be the right and left *Perron vectors* of $\mathbf{A}$, the eigenvectors associated with the Perron root, normalized so that $\mathbf{e}^\top \mathbf{v}(\mathbf{A}) = \mathbf{u}(\mathbf{A})^\top \mathbf{v}(\mathbf{A}) = 1$, where $\mathbf{e}$ is the vector of ones. Since $\mathbf{A}$ is irreducible, from Perron-Frobenius theory, $\mathbf{v}(\mathbf{A})$ and $\mathbf{u}(\mathbf{A})^\top$ are strictly positive and unique.

The goal is to provide conditions that make the spectral radius of $\mathbf{A}[(1-m)\, r_B\, \mathbf{I} + m\mathbf{B}]\mathbf{D}$ monotonic in m. We proceed as follows: first, $\mathbf{A}$ and $\mathbf{B}$ are constrained to commute and be symmetrizable, which allows them to be simultaneously represented by the canonical form (6); second, this form is used to show that its spectral radius can be represented as a sum of squares; finally, the

derivative of the spectral radius is represented as a sum of squares, and this is utilized to give conditions that determine its sign.

The following representation of symmetrizable matrices is used throughout. It arises for the special case of transition matrices of reversible Markov chains [16, p. 33]:

Lemma 4 (Canonical Form for Symmetrizable Matrices [2, Lemma 1])**.** *A symmetrizable matrix* $\mathbf{A} = \mathbf{D}_\mathsf{L}\mathbf{S}\mathbf{D}_\mathsf{R}$*, where* $\mathbf{S}$ *is symmetric and* $\mathbf{D}_\mathsf{L}$ *and* $\mathbf{D}_\mathsf{R}$ *are positive diagonal matrices, can always be put into a canonical form*

$$\mathbf{A} = \mathbf{D}_\mathsf{L}\mathbf{S}\mathbf{D}_\mathsf{R} = \mathbf{E}\mathbf{K}\mathbf{\Lambda}\mathbf{K}^\top\mathbf{E}^{-1}, \tag{6}$$

where $\mathbf{E} = \mathbf{D}_\mathsf{L}^{1/2}\mathbf{D}_\mathsf{R}^{-1/2}$ *is a positive diagonal matrix,* $\mathbf{K}$ *is an orthogonal matrix,* $\mathbf{K}^\top$ *is its transpose,* $\mathbf{\Lambda}$ *is a diagonal matrix of the eigenvalues of* $\mathbf{A}$*, the columns of* $\mathbf{EK}$ *are right eigenvectors of* $\mathbf{A}$*, and the rows of* $\mathbf{K}^\top\mathbf{E}^{-1}$ *are left eigenvectors of* $\mathbf{A}$*.*

Proof. First it is shown that for arbitrary symmetric $\mathbf{S}$ there is a symmetric $\hat{\mathbf{S}}$ such that $\mathbf{A} = \mathbf{D}_\mathsf{L}\mathbf{S}\mathbf{D}_\mathsf{R} = \mathbf{E}\hat{\mathbf{S}}\mathbf{E}^{-1}$. Direct substitution of $\mathbf{E} = \mathbf{D}_\mathsf{L}^{1/2}\mathbf{D}_\mathsf{R}^{-1/2}$ gives

$$\begin{aligned}
&\mathbf{D}_\mathsf{L}\mathbf{S}\mathbf{D}_\mathsf{R} = \mathbf{D}_\mathsf{L}^{1/2}\mathbf{D}_\mathsf{R}^{-1/2}\hat{\mathbf{S}}\mathbf{D}_\mathsf{L}^{-1/2}\mathbf{D}_\mathsf{R}^{1/2}\\
&\iff \hat{\mathbf{S}} = \mathbf{D}_\mathsf{L}^{-1/2}\mathbf{D}_\mathsf{R}^{1/2}\mathbf{D}_\mathsf{L}\mathbf{S}\mathbf{D}_\mathsf{R}\mathbf{D}_\mathsf{L}^{1/2}\mathbf{D}_\mathsf{R}^{-1/2} = \mathbf{D}_\mathsf{L}^{1/2}\mathbf{D}_\mathsf{R}^{1/2}\mathbf{S}\mathbf{D}_\mathsf{L}^{1/2}\mathbf{D}_\mathsf{R}^{1/2},
\end{aligned}$$

which is symmetric. Symmetric $\hat{\mathbf{S}}$ has a symmetric Jordan canonical form $\hat{\mathbf{S}} = \mathbf{K}\mathbf{\Lambda}\mathbf{K}^\top$ where $\mathbf{K}$ is an orthogonal matrix and $\mathbf{\Lambda}$ is a matrix of the eigenvalues of $\hat{\mathbf{S}}$, which by construction are also the eigenvalues of $\mathbf{A}$. Hence $\mathbf{A} = \mathbf{D}_\mathsf{L}\mathbf{S}\mathbf{D}_\mathsf{R} = \mathbf{E}\hat{\mathbf{S}}\mathbf{E}^{-1} = \mathbf{E}\mathbf{K}\mathbf{\Lambda}\mathbf{K}^\top\mathbf{E}^{-1}$. Let $[\mathbf{EK}]_i$ be the ith column of $\mathbf{EK}$. Then

$$\mathbf{A}[\mathbf{EK}]_i = \mathbf{E}\mathbf{K}\mathbf{\Lambda}\mathbf{K}^\top\mathbf{E}^{-1}[\mathbf{EK}]_i = \mathbf{E}\mathbf{K}\mathbf{\Lambda}\mathbf{K}^\top[\mathbf{K}]_i = \mathbf{E}\mathbf{K}\mathbf{\Lambda}\mathbf{e}_i = \lambda_i\mathbf{E}\mathbf{K}\mathbf{e}_i = \lambda_i[\mathbf{EK}]_i$$

hence $[\mathbf{EK}]_i$ is a right eigenvector of $\mathbf{A}$. The analogous derivation shows the rows of $\mathbf{K}^\top\mathbf{E}^{-1}$ to be left eigenvectors of $\mathbf{A}$. ☐

Lemma 5 (Canonical Form for Commuting Symmetrizable $\mathbf{A}$ and $\mathbf{B}$)**.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be* $n \times n$ *symmetrizable irreducible nonnegative matrices that commute with each other. Then* $\mathbf{A}$ *and* $\mathbf{B}$ *can be decomposed as*

$$\mathbf{A} = \mathbf{D}_\mathbf{v}^{1/2}\mathbf{D}_\mathbf{u}^{-1/2}\mathbf{K}\mathbf{\Lambda}_A\mathbf{K}^\top\mathbf{D}_\mathbf{v}^{-1/2}\mathbf{D}_\mathbf{u}^{1/2}, \tag{7}$$

$$\mathbf{B} = \mathbf{D}_\mathbf{v}^{1/2}\mathbf{D}_\mathbf{u}^{-1/2}\mathbf{K}\mathbf{\Lambda}_B\mathbf{K}^\top\mathbf{D}_\mathbf{v}^{-1/2}\mathbf{D}_\mathbf{u}^{1/2}, \tag{8}$$

where $\mathbf{v} \equiv \mathbf{v}(\mathbf{A}) = \mathbf{v}(\mathbf{B})$*,* $\mathbf{u} \equiv \mathbf{u}(\mathbf{A}) = \mathbf{u}(\mathbf{B})$*,* $\mathbf{K}$ *is an orthogonal matrix,* $[\mathbf{K}]_1 = \mathbf{D}_\mathbf{v}^{1/2}\mathbf{D}_\mathbf{u}^{1/2}\mathbf{e}$*, and* $\mathbf{\Lambda}_A$ *and* $\mathbf{\Lambda}_B$ *are diagonal matrices of the eigenvalues of* $\mathbf{A}$ *and* $\mathbf{B}$*, respectively.*

Proof. Since $\mathbf{A}$ and $\mathbf{B}$ are symmetrizable, each can be represented by canonical form (6). $\mathbf{A}$ and $\mathbf{B}$ are diagonalizable since $\mathbf{\Lambda}$ in (6) is a diagonal matrix and $(\mathbf{EK})^{-1} = \mathbf{K}^\top\mathbf{E}^{-1}$. Since $\mathbf{A}$ and $\mathbf{B}$ commute by hypothesis, they can be

simultaneously diagonalized [14, Theorem 1.3.19, p. 52], which means there exists an invertible $\mathbf{X}$ such that $\mathbf{A} = \mathbf{X}\mathbf{\Lambda}_A\mathbf{X}^{-1}$ and $\mathbf{B} = \mathbf{X}\mathbf{\Lambda}_B\mathbf{X}^{-1}$. Clearly the columns of $\mathbf{X}$ are right eigenvectors of $\mathbf{A}$ and $\mathbf{B}$, and the rows of $\mathbf{X}^{-1}$ are left eigenvectors of $\mathbf{A}$ and $\mathbf{B}$, since

$$\mathbf{A}[\mathbf{X}]_i = \mathbf{X}\mathbf{\Lambda}_A\mathbf{X}^{-1}[\mathbf{X}]_i = \mathbf{X}\mathbf{\Lambda}_A\mathbf{e}_i = \mathbf{X}\lambda_i(\mathbf{A})\mathbf{e}_i = \lambda_i(\mathbf{A})[\mathbf{X}]_i,$$

etc.. We can set $\mathbf{X} = \mathbf{EK}$ to give

$$\begin{aligned}\mathbf{A} &= \mathbf{X}\mathbf{\Lambda}_A\mathbf{X}^{-1} = \mathbf{EK}\mathbf{\Lambda}_A(\mathbf{EK})^{-1} = \mathbf{EK}\mathbf{\Lambda}_A\mathbf{K}^\top\mathbf{E}^{-1}\\ \mathbf{B} &= \mathbf{X}\mathbf{\Lambda}_B\mathbf{X}^{-1} = \mathbf{EK}\mathbf{\Lambda}_B(\mathbf{EK})^{-1} = \mathbf{EK}\mathbf{\Lambda}_B\mathbf{K}^\top\mathbf{E}^{-1}.\end{aligned}$$

Without loss of generality, the Perron root is given index 1, so $r(\mathbf{A}) = \lambda_1(\mathbf{A})$, $r(\mathbf{B}) = \lambda_1(\mathbf{B})$. Since $\mathbf{A}$ and $\mathbf{B}$ are irreducible,

$$\mathbf{v}(\mathbf{A}) = [\mathbf{EK}]_1 = \mathbf{v}(\mathbf{B}) \equiv \mathbf{v} > \mathbf{0}, \tag{9}$$

$$\mathbf{u}(\mathbf{A})^\top = [\mathbf{K}^\top\mathbf{E}^{-1}]^1 = \mathbf{u}(\mathbf{B})^\top \equiv \mathbf{u}^\top > \mathbf{0}. \tag{10}$$

Next, $\mathbf{E}$ is solved in terms of $\mathbf{u}^\top$ and $\mathbf{v}$:

$$[\mathbf{EK}]_1 = \mathbf{E}[\mathbf{K}]_1 = \mathbf{v}, \text{ so } [\mathbf{K}]_1 = \mathbf{E}^{-1}\mathbf{v},$$

and

$$[\mathbf{K}^\top\mathbf{E}^{-1}]^1 = [\mathbf{K}^\top]^1\mathbf{E}^{-1} = \mathbf{u}^\top \text{ so } [\mathbf{K}^\top]^1 = \mathbf{u}^\top\mathbf{E},$$

which combined give $K_{1j} = E_j^{-1}v_j = u_jE_j$, hence $E_j^2 = v_j/u_j$ so $E_j = \sqrt{v_j/u_j}$ and

$$\mathbf{E} = \mathbf{D}_\mathbf{v}^{1/2}\mathbf{D}_\mathbf{u}^{-1/2}. \tag{11}$$

The first column of $\mathbf{K}$ evaluates to

$$[\mathbf{K}]_1 = \mathbf{E}^{-1}\mathbf{v} = \mathbf{D}_\mathbf{v}^{-1/2}\mathbf{D}_\mathbf{u}^{1/2}\mathbf{v} = \mathbf{D}_\mathbf{v}^{1/2}\mathbf{D}_\mathbf{u}^{1/2}\mathbf{e}. \tag{12}$$

□

Theorem 6 (Sum-of-Squares Solution for the Spectral Radius)**.**
Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ symmetrizable irreducible nonnegative matrices that commute. Let $r_A \equiv r(\mathbf{A}) = \lambda_{A1}$ and $r_B \equiv r(\mathbf{B}) = \lambda_{B1}$ refer to their Perron roots, and $\{\lambda_{Ai}\}$ and $\{\lambda_{Bi}\}$ represent all of the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$, respectively. Let $\mathbf{u}^\top$ and $\mathbf{v}$ be the common left and right Perron vectors of $\mathbf{A}$ and $\mathbf{B}$ (Lemma 5). Let $\mathbf{D}$ be a positive diagonal matrix and define

$$\mathbf{M}(m) := \mathbf{A}[(1-m)\,r_B\,\mathbf{I} + m\mathbf{B}], \qquad m \in [0,1].$$

Let $\mathbf{v}(m) \equiv \mathbf{v}(\mathbf{M}(m)\mathbf{D})$ and $\mathbf{u}(m)^\top \equiv \mathbf{u}(\mathbf{M}(m)\mathbf{D})^\top$ refer to the right and left Perron vectors of $\mathbf{M}(m)\mathbf{D}$.

Then

$$r(\mathbf{M}(m)\mathbf{D}) = \sum_{i=1}^{n} \lambda_{Ai}\,[(1-m)r_B + m\lambda_{Bi}]\,y_i(m)^2, \tag{13}$$

where

$$y_i(m)^2 = \frac{1}{\displaystyle\sum_j v_j(m)^2 \frac{D_j u_j}{v_j}} \left(\sum_j K_{ji} \frac{D_j u_j^{1/2} v_j(m)}{v_j^{1/2}} \right)^2. \tag{14}$$

and $\mathbf{K}$ *is from the canonical form in Lemma 5.*

Proof. One can represent $\mathbf{M}(m)$ using the canonical forms (7), (8):

$$\mathbf{M}(m) = \mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{K}\mathbf{\Lambda}_A[(1-m)\,r_B\,\mathbf{I} + m\mathbf{\Lambda}_B]\mathbf{K}^\top\mathbf{D}_{\mathbf{v}}^{-1/2}\mathbf{D}_{\mathbf{u}}^{1/2}. \tag{15}$$

This form will be used to produce a symmetric matrix similar to $\mathbf{M}(m)\mathbf{D}$, which allows use of the Rayleigh-Ritz variational formula for the spectral radius. The expression will be seen to simplify to the sum of squared terms.

For brevity let $\mathbf{\Phi}_m := \mathbf{K}\mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]\mathbf{K}^\top$, so $\mathbf{M}(m) = \mathbf{E}\mathbf{\Phi}_m\mathbf{E}^{-1}$ where $\mathbf{E} = \mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}$ from (11). Since $\mathbf{M}(m) \geq \mathbf{0}$ and $\mathbf{E}$ is a positive diagonal matrix, then $\mathbf{\Phi}_m \geq \mathbf{0}$. The following identities are obtained by multiplication on both sides by $\mathbf{E}$, $\mathbf{D}^{1/2}$, and their inverses:

$$\begin{aligned} r(\mathbf{M}(m)\mathbf{D}) &= r(\mathbf{E}\mathbf{\Phi}_m\mathbf{E}^{-1}\mathbf{D}) = r(\mathbf{\Phi}_m\mathbf{E}^{-1}\mathbf{D}\mathbf{E}) = r(\mathbf{\Phi}_m\mathbf{D}) = r(\mathbf{D}^{1/2}\mathbf{\Phi}_m\mathbf{D}^{1/2}) \\ &= r(\mathbf{S}_m), \end{aligned}$$

where

$$\begin{aligned} \mathbf{S}_m &:= \mathbf{D}^{1/2}\mathbf{\Phi}_m\mathbf{D}^{1/2} = \mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]\mathbf{K}^\top\mathbf{D}^{1/2} \qquad (16) \\ &= \mathbf{D}^{1/2}\mathbf{E}^{-1}\mathbf{M}(m)\mathbf{E}\mathbf{D}^{1/2} \qquad (17) \end{aligned}$$

is symmetric, so we may apply the Rayleigh-Ritz variational formula for the spectral radius [14, Theorem 4.2.2, p. 176]:

$$r(\mathbf{S}_m) = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\mathbf{x}^\top\mathbf{S}_m\mathbf{x}}{\mathbf{x}^\top\mathbf{x}}. \tag{18}$$

This yields

$$\begin{aligned} r(\mathbf{M}(m)\mathbf{D}) = r(\mathbf{S}_m) &= \max_{\mathbf{x}^\top\mathbf{x}=1} \mathbf{x}^\top\mathbf{S}_m\mathbf{x} \\ &= \max_{\mathbf{x}^\top\mathbf{x}=1} \mathbf{x}^\top\mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]\mathbf{K}^\top\mathbf{D}^{1/2}\mathbf{x}. \qquad (19) \end{aligned}$$

Any $\hat{\mathbf{x}}$ that yields the maximum in (19) is an eigenvector of $\mathbf{S}_m$ [13, p. 33]. Because $\mathbf{M}(m)$ is irreducible and $\mathbf{D}$ and $\mathbf{E}$ are positive diagonal matrices, $\mathbf{M}(m)\mathbf{D}$

and $\mathbf{S}_m = \mathbf{D}^{1/2}\mathbf{E}^{-1}\mathbf{M}(m)\mathbf{E}\mathbf{D}^{1/2}$ are irreducible. By Perron-Frobenius theory, $\hat{\mathbf{x}}(m) > 0$ is therefore the unique left and right Perron vector of $\mathbf{S}_m$. This allows one to write

$$r(\mathbf{M}(m)\mathbf{D}) = \hat{\mathbf{x}}(m)^\top \mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]\mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m). \tag{20}$$

Define

$$\mathbf{y}(m) := \mathbf{K}^\top \mathbf{D}^{1/2}\hat{\mathbf{x}}(m). \tag{21}$$

Substitution of (21) into (20) yields (13):

$$\begin{aligned} r(\mathbf{M}(m)\mathbf{D}) &= \hat{\mathbf{x}}(m)^\top \mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]\mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m) \\ &= \mathbf{y}(m)^\top \mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]\,\mathbf{y}(m) \\ &= \sum_{i=1}^n \lambda_{Ai}[(1-m)r_B + m\lambda_{Bi}]\,y_i(m)^2. \end{aligned}$$

Next, $\mathbf{y}(m)$ will be solved in terms of $\mathbf{v}(m)$ and $\mathbf{u}(m)$ by solving for $\hat{\mathbf{x}}(m)$, using the following two facts. For brevity, define $\mathbf{\Lambda}_m := \mathbf{\Lambda}_A[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]$, so $\mathbf{M}(m) = \mathbf{E}\mathbf{K}\mathbf{\Lambda}_m\mathbf{K}^\top\mathbf{E}^{-1}$:

1. $r(\mathbf{M}(m)\mathbf{D})\ \mathbf{v}(m) = \mathbf{M}(m)\mathbf{D}\,\mathbf{v}(m) = \mathbf{E}\mathbf{K}\mathbf{\Lambda}_m\mathbf{K}^\top\mathbf{E}^{-1}\mathbf{D}\,\mathbf{v}(m);$ (22)
2. $r(\mathbf{M}(m)\mathbf{D})\ \hat{\mathbf{x}}(m) = \mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_m\mathbf{K}^\top\mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m).$ (23)

Multiplication on the left by $\mathbf{E}\mathbf{D}^{-1/2}$ on both sides of (23) reveals the right Perron vector $\mathbf{v}(m) = \mathbf{v}(\mathbf{M}(m)\mathbf{D})$:

$$\begin{aligned} r(\mathbf{M}(m)\mathbf{D})\ (\mathbf{E}\mathbf{D}^{-1/2})\hat{\mathbf{x}}(m) &= \ (\mathbf{E}\mathbf{D}^{-1/2})\mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_m\mathbf{K}^\top\mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m) \\ &= \mathbf{E}\mathbf{K}\mathbf{\Lambda}_m\mathbf{K}^\top(\mathbf{E}^{-1}\mathbf{D}\mathbf{E}\mathbf{D}^{-1})\mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m) \\ &= (\mathbf{E}\mathbf{K}\mathbf{\Lambda}_m\mathbf{K}^\top\mathbf{E}^{-1}\mathbf{D})(\mathbf{E}\mathbf{D}^{-1/2}\hat{\mathbf{x}}(m)) \\ &= \mathbf{M}(m)\mathbf{D}(\mathbf{E}\mathbf{D}^{-1/2}\hat{\mathbf{x}}(m)), \end{aligned} \tag{24}$$

which shows that $\mathbf{E}\mathbf{D}^{-1/2}\hat{\mathbf{x}}(m)$ is the right Perron vector of $\mathbf{M}(m)\mathbf{D}$, unique up to scaling, i.e.

$$\mathbf{v}(m) = \hat{c}(m)\,\mathbf{E}\mathbf{D}^{-1/2}\hat{\mathbf{x}}(m) = \hat{c}(m)\,\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{D}^{-1/2}\hat{\mathbf{x}}(m),$$

for some $\hat{c}(m)$ to be solved, which gives

$$\hat{\mathbf{x}}(m) = \frac{1}{\hat{c}(m)}\mathbf{D}_{\mathbf{v}}^{-1/2}\mathbf{D}_{\mathbf{u}}^{1/2}\mathbf{D}^{1/2}\ \mathbf{v}(m). \tag{25}$$

The normalization constraint $\hat{\mathbf{x}}(m)^\top\hat{\mathbf{x}}(m) = 1$ yields

$$1 = \hat{\mathbf{x}}(m)^\top\hat{\mathbf{x}}(m) = \frac{1}{\hat{c}(m)^2}\mathbf{v}(m)^\top\mathbf{D}_{\mathbf{v}}^{-1}\mathbf{D}_{\mathbf{u}}\mathbf{D}\,\mathbf{v}(m),$$

so

$$\hat{c}(m) = \sqrt{\mathbf{v}(m)^\top\, \mathbf{D}_\mathbf{v}^{-1}\mathbf{D}_\mathbf{u}\mathbf{D}\,\mathbf{v}(m)}. \tag{26}$$

Substitution for $\hat{\mathbf{x}}(m)$ now produces (14):

$$\begin{aligned}\mathbf{y}(m) &:= \mathbf{K}^\top\mathbf{D}^{1/2}\hat{\mathbf{x}}(m) = \mathbf{K}^\top\mathbf{D}^{1/2}\frac{1}{\hat{c}(m)}\mathbf{D}_\mathbf{v}^{-1/2}\mathbf{D}_\mathbf{u}^{1/2}\mathbf{D}^{1/2}\ \mathbf{v}(m)\\ &= \frac{1}{\sqrt{\mathbf{v}(m)^\top\ \mathbf{D}_\mathbf{v}^{-1}\mathbf{D}_\mathbf{u}\mathbf{D}\,\mathbf{v}(m)}}\ \mathbf{K}^\top\mathbf{D}_\mathbf{v}^{-1/2}\mathbf{D}_\mathbf{u}^{1/2}\mathbf{D}\,\mathbf{v}(m). \end{aligned}\tag{27}$$

Each element of $\mathbf{y}(m)$ is thus

$$y_i(m) = \frac{1}{\sqrt{\sum_j v_j(m)^2\frac{D_j u_j}{v_j}}}\sum_j K_{ji}\frac{D_j u_j^{1/2} v_j(m)}{v_j^{1/2}},$$

hence

$$y_i(m)^2 = \frac{1}{\sum_j v_j(m)^2\frac{D_j u_j}{v_j}}\left(\sum_j K_{ji}\frac{D_j u_j^{1/2} v_j(m)}{v_j^{1/2}}\right)^2.$$

□

Theorem 7 (Main Result)**.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be* $n\times n$ *symmetrizable irreducible nonnegative matrices that commute with each other, with Perron roots* $r_A \equiv r(\mathbf{A}) = \lambda_{A1}$ *and* $r_B \equiv r(\mathbf{B}) = \lambda_{B1}$*, and common left and right Perron vectors,* $\mathbf{u}^\top$ *and* $\mathbf{v}$*. Let* $\mathbf{D}$ *be a nonscalar positive diagonal matrix, and suppose*

$$\mathbf{M}(m) := \mathbf{A}[(1-m)\, r_B\,\mathbf{I} + m\mathbf{B}], \qquad m \in [0,1]. \tag{28}$$

C1. If all eigenvalues of $\mathbf{A}$ *are positive, then*

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) < 0. \tag{29}$$

C2. If all eigenvalues of $\mathbf{A}$ *other than* $r_A = \lambda_1(\mathbf{A})$ *are negative, then*

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) > 0. \tag{30}$$

C3. If $\lambda_{Ai} = 0$ *for all* $i \in \{2,\dots,n\}$*, then* $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m = 0$.

C4. If C1 or C2 hold except for some $i \in \{2,\dots,n\}$ *where* $\lambda_{Ai} = 0$*, then inequalities* (29) *and* (30) *are replaced by non-strict inequalities.*

Proof. The sum-of-squares form in Theorem 6 is now utilized to analyze the derivative of the spectral radius. For an arbitrary irreducible nonnegative matrix $\mathbf{F}(m)$ that is a differentiable function of m, the derivative of its spectral radius follows the general relation [8, Sec. 9.1.1]:

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{F}(m)) = \mathbf{u}(\mathbf{F}(m))^\top \frac{\mathrm{d}\mathbf{F}(m)}{\mathrm{d}m} \mathbf{v}(\mathbf{F}(m)). \tag{31}$$

The derivatives of $\mathbf{u}(\mathbf{F}(m))$ and $\mathbf{v}(\mathbf{F}(m))$ do not appear in (31) because they are critical points with respect to $r(\mathbf{F}(m))$. Application of (31) gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) &= \hat{\mathbf{x}}(m)^\top \mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_A \frac{\mathrm{d}[(1-m)r_B\mathbf{I} + m\mathbf{\Lambda}_B]}{\mathrm{d}m} \mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m)\\ &= \hat{\mathbf{x}}(m)^\top \mathbf{D}^{1/2}\mathbf{K}\mathbf{\Lambda}_A(\mathbf{\Lambda}_B - r_B\mathbf{I})\mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m).\end{aligned}$$

Substitution with $\mathbf{y}(m) := \mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m)$ yields:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) &= \mathbf{y}(m)^\top \mathbf{\Lambda}_A(\mathbf{\Lambda}_B - r_B\mathbf{I})\,\mathbf{y}(m)\\ &= \sum_{i=1}^{n} \lambda_{Ai}\,(\lambda_{Bi} - r_B)\,y_i(m)^2. \end{aligned}\tag{32}$$

We know the following about the terms in the sum in (32):

1. $\lambda_{B1} - r_B = 0$. Thus the first term $i = 1$ of the sum is zero.
2. *For* $i \in \{2, \ldots, n\}$, $\lambda_{Bi} - r_B < 0$, *hence* $(\lambda_{Bi} - r_B)y_i^2 \leq 0$. Since $\mathbf{B}$ is symmetrizable, $\lambda_{Bi} \in \mathbb{R}$. Since $\mathbf{B}$ is irreducible the Perron root has multiplicity 1, and $|\lambda_{Bi}| \leq r_B$ [21, Theorems 1.1, 1.5]. Together these imply $\lambda_{Bi} < r_B$ for $i \in \{2, \ldots, n\}$.
3. $y_i(m) \neq 0$ *for at least one* $i \in \{2, \ldots, n\}$, *whenever* $\mathbf{D} \neq c\,\mathbf{I}$ *for any* $c > 0$. Suppose to the contrary that $y_i(m) = 0$ for all $i \in \{2, \ldots, n\}$. That means $\mathbf{y}(m) = y_1(m)\,\mathbf{e}_1$ so (27) becomes

$$\mathbf{y}(m) = \; y_1(m)\,\mathbf{e}_1 = \hat{c}(m)^{-1}\;\mathbf{K}^\top \mathbf{D}_{\mathbf{v}}^{-1/2}\mathbf{D}_{\mathbf{u}}^{1/2}\mathbf{D}\,\mathbf{v}.$$

   Now multiply on the left by nonsingular $\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{K}$ and substitute $[\mathbf{K}]_1 = \mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{1/2}\mathbf{e}$ (12):

$$\begin{aligned}y_1(m)\,(\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{K})\,\mathbf{e}_1 &= y_1(m)\,\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}[\mathbf{K}]_1\\ &= y_1(m)\,\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{1/2}\mathbf{e}\\ &= y_1(m)\,\mathbf{v}\\ = \hat{c}(m)^{-1}(\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{K})\mathbf{K}^\top \mathbf{D}_{\mathbf{v}}^{-1/2}\mathbf{D}_{\mathbf{u}}^{1/2}\mathbf{D}\,\mathbf{v}\\ &= \hat{c}(m)^{-1}\;\mathbf{D}\,\mathbf{v}.\end{aligned}$$

   Hence $\mathbf{v} = (y_1(m)/\hat{c}(m))\mathbf{D}\mathbf{v} > \mathbf{0}$, which implies $\mathbf{D} = (\hat{c}(m)/y_1(m))\mathbf{I}$, contrary to hypothesis that $\mathbf{D} \neq c\,\mathbf{I}$ for any $c > 0$. Thus $\mathbf{D}$ being nonscalar implies that $y_i(m) \neq 0$ for at least one $i \in \{2, \ldots, n\}$.

Combining points 2., and 3. above, we have $(\lambda_{Bi}-r_B)y_i(m)^2 < 0$ for at least one $i \in \{2, \ldots, n\}$, while from point 1., $\lambda_{A1}(\lambda_{B1} - r_B)y_1(m)^2 = 0$. Thus, if the signs of λ_{Ai}, $i = 2, \ldots, n$ are the same, the nonzero terms in the sum in (32) all have the same, opposite sign, and there is at least one such nonzero term. Therefore,

1. if $\lambda_{Ai} > 0$ for all i, then $\lambda_{Ai}(\lambda_{Bi} - r_B) < 0\ \forall i \geq 2$, thus

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) = \sum_{i=2}^{n} \lambda_{Ai}(\lambda_{Bi} - r_B)y_i^2(m) < 0;$$

2. if $\lambda_{Ai} < 0$ for $i = 2, \ldots, n$, then $\lambda_{Ai}(\lambda_{Bi} - r_B) > 0\ \forall i \geq 2$, thus

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) = \sum_{i=2}^{n} \lambda_{Ai}(\lambda_{Bi} - r_B)y_i^2(m) > 0;$$

3. If $\lambda_{Ai} = 0$ for all $i \in \{2, \ldots, n\}$, then all the terms in (32) are zero so $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m = 0$.

4. If $\lambda_{Ai} = 0$ for some $i \in \{2, \ldots, n\}$, we cannot exclude the possibility that $y_i(m)$ happens to be the one necessary nonzero value among $y_2(m), \ldots, y_n(m)$, while $y_j(m) = 0$ for all $j \neq 1, i$, in which case all the terms in (32) would be zero. Thus inequalities in (29) and (30) become non-strict if any $\lambda_{Ai} = 0$.

If the non-Perron eigenvalues of $\mathbf{A}$ are a mix of positive, negative, or zero values, there may be positive, negative, or zero terms $\lambda_{Ai}(\lambda_{Bi} - r_B)y_i^2$ for $i = 2, \ldots, n$, so the sign of $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m$ depends on the particular magnitudes of λ_{Ai}, λ_{Bi}, r_B, and $y_i(m)$.

□

Corollary 8. *For the case $n = 2$ of Theorem 7, $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m$ has the opposite sign of λ_{A2} and is zero if $\lambda_{A2} = 0$.*

Proof. When $n = 2$, $\mathbf{A}$ has only one eigenvalue besides the Perron root. Therefore $\mathbf{D} \neq c\,\mathbf{I}$ implies $y_2(m) \neq 0$ in (32), so $(\lambda_{B2} - r_B)y_2^2 < 0$. Thus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) &= \lambda_{A1}(\lambda_{B1} - r_B)y_1^2 + \lambda_{A2}(\lambda_{B2} - r_B)y_2^2 \\ &= r_A * 0 * y_1^2 + \lambda_{A2}(\lambda_{B2} - r_B)y_2^2.\end{aligned}$$

Therefore $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m$ has the opposite sign of λ_{A2}, or is 0 if $\lambda_{A2} = 0$. □

The following is immediate:

Corollary 9. *In Theorem 7, the term $r_B\mathbf{I}$ in (28) may be replaced by any symmetrizable nonnegative matrix $\mathbf{C}$ that commutes with $\mathbf{A}$ and $\mathbf{B}$ for which $r(\mathbf{C}) = r(\mathbf{B})$ and $\lambda_{Ci} > \lambda_{Bi}$, $i = 2, \ldots, n$.*

Note that the indices $i \in \{2, \ldots, n\}$ are not ordered here by the size of the eigenvalues as is commonly done, but are set by the arbitrary indexing of the non-Perron eigenvectors.

Corollary 10 (Partial answer to Cohen's open question)**.** *Let $\mathbf{A}$ be an $n \times n$ symmetrizable nonnegative matrix and $\mathbf{D}$ be a positive diagonal matrix. If all of the eigenvalues of $\mathbf{A}$ are positive, then $r(\mathbf{A})\, r(\mathbf{AD}) \geq r(\mathbf{A}^2\mathbf{D})$. If all of the non-Perron eigenvalues of $\mathbf{A}$ are negative, then $r(\mathbf{A})\, r(\mathbf{AD}) \leq r(\mathbf{A}^2\mathbf{D})$. When $\mathbf{A}$ is irreducible and $\mathbf{D}$ is nonscalar, then the above inequalities are strict. If all of the non-Perron eigenvalues of $\mathbf{A}$ are zero, then $r(\mathbf{A})\, r(\mathbf{AD}) = r(\mathbf{A}^2\mathbf{D})$.*

Proof. Let $\mathbf{A}$ be irreducible and $\mathbf{D}$ nonscalar. Apply Theorem 7 with $\mathbf{A} = \mathbf{B}$. Then $\mathbf{M}(0) = r(\mathbf{A})\mathbf{A}$ and $\mathbf{M}(1) = \mathbf{A}^2$. If all the eigenvalues of $\mathbf{A}$ are positive, then by Theorem 7, $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m < 0$, so $r(\mathbf{M}(0)\mathbf{D}) = r(\mathbf{A})r(\mathbf{AD}) > r(\mathbf{A}^2\mathbf{D}) = r(\mathbf{M}(1)\mathbf{D})$. If all the non-Perron eigenvalues of $\mathbf{A}$ are negative, then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m > 0$, so $r(\mathbf{M}(0)\mathbf{D}) = r(\mathbf{A})r(\mathbf{AD}) < r(\mathbf{A}^2\mathbf{D}) = r(\mathbf{M}(1)\mathbf{D})$. If $\lambda_{Ai} = 0$ for $i \in \{2, \ldots, n\}$ then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m = 0$ so $r(\mathbf{A})r(\mathbf{AD}) = r(\mathbf{A}^2\mathbf{D})$.

If $\mathbf{D} = c\,\mathbf{I}$ for $c > 0$ then $r(\mathbf{A})\, r(\mathbf{AD}) = c\, r(\mathbf{A})^2 = c\, r(\mathbf{A}^2)$ so equality holds. A reducible symmetrizable nonnegative matrix $\mathbf{A}$ is always the limit of some sequence of symmetrizable irreducible nonnegative matrices, for which the eigenvalues remain on the real line. If $\lambda_{Ai}, i = 2, \ldots, n$ are all negative or all positive, then they continue to be so for these perturbations of $\mathbf{A}$ by the continuity of the eigenvalues. For each perturbation, the sign of $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m$ is maintained, but in the limit equality cannot be excluded, so only the non-strict versions of the inequalities are assured for reducible matrices. □

Theorem 11 (Main Result Variant)**.** *Let $\mathbf{A}$ and $\mathbf{B}$ be $n \times n$ symmetrizable irreducible nonnegative matrices that commute with each other, with equal Perron roots $r(\mathbf{A}) = r(\mathbf{B}) = \lambda_{A1} = \lambda_{B1}$, and common left and right Perron vectors, $\mathbf{u}^\top$ and $\mathbf{v}$. Let λ_{Ai} and λ_{Bi}, $i = 2, \ldots, n$ be the non-Perron eigenvalues. Let $\mathbf{D}$ be a nonscalar positive diagonal matrix, and suppose*

$$\mathbf{M}(m) := (1-m)\mathbf{A} + m\mathbf{B}, \qquad m \in [0, 1]. \tag{33}$$

1. *If $\lambda_{Ai} > \lambda_{Bi}$ for $i = 2, \ldots n$, then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m < 0$ and $r(\mathbf{AD}) > r(\mathbf{BD})$.*
2. *If $\lambda_{Ai} < \lambda_{Bi}$ for $i = 2, \ldots n$, then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m > 0$ and $r(\mathbf{AD}) < r(\mathbf{BD})$.*
3. *If $\lambda_{Ai} = \lambda_{Bi}$ for $i = 2, \ldots n$, then $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m = 0$ and $r(\mathbf{AD}) = r(\mathbf{BD})$.*
4. *If $\lambda_{Ai} = \lambda_{Bi}$ for at least one $i \in \{2, \ldots, n\}$, then the inequalities in 1 and 2 are replaced by non-strict inequalities.*

Proof. The proof follows that of Theorem 7 with some substitutions. $\mathbf{M}(m)$ has the canonical representation

$$\mathbf{M}(m) := (1-m)\mathbf{A} + m\mathbf{B} = \mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{D}_{\mathbf{u}}^{-1/2}\mathbf{K}[(1-m)\,\mathbf{\Lambda}_A + m\mathbf{\Lambda}_B]\mathbf{K}^\top\mathbf{D}_{\mathbf{v}}^{-1/2}\mathbf{D}_{\mathbf{u}}^{1/2}.$$

The spectral radius has the sum-of-squares form as developed in (16)–(27), where $\hat{\mathbf{x}}(m)$ is as given in (25), and the derivative of the spectral radius evaluates to

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) &= \hat{\mathbf{x}}(m)^\top \mathbf{D}^{1/2}\mathbf{K}\frac{\mathrm{d}[(1-m)\mathbf{\Lambda}_A + m\mathbf{\Lambda}_B]}{\mathrm{d}m}\mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m)\\ &= \hat{\mathbf{x}}(m)^\top \mathbf{D}^{1/2}\mathbf{K}(\mathbf{\Lambda}_B - \mathbf{\Lambda}_A)\mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m).\end{aligned}$$

Substitution with $\mathbf{y}(m) := \mathbf{K}^\top \mathbf{D}^{1/2}\,\hat{\mathbf{x}}(m)$ yields:

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{M}(m)\mathbf{D}) = \mathbf{y}(m)^\top (\mathbf{\Lambda}_B - \mathbf{\Lambda}_A)\,\mathbf{y}(m) = \sum_{i=1}^{n} (\lambda_{Bi} - \lambda_{Ai})\, y_i(m)^2. \tag{34}$$

The relevant facts about (34) are:

1. $\lambda_{B1} - \lambda_{Ai} = 0$ by construction. Thus the first term $i = 1$ of the sum is zero.
2. $y_i(m) \neq 0$ for at least one $i \in \{2, \dots, n\}$, whenever $\mathbf{D} \neq c\,\mathbf{I}$ for any $c > 0$, as in Theorem 7.

If $\lambda_{Ai} > \lambda_{Bi}$ for $i = 2, \dots, n$ then all of the terms $(\lambda_{Bi} - \lambda_{Ai})\, y_i(m)^2$ in (34) are nonpositive, and at least one is negative, therefore $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m$ is negative. If $\lambda_{Ai} < \lambda_{Bi}$ for $i = 2, \dots, n$ then all of the terms in (34) are nonnegative, and at least one is positive, therefore $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m$ is positive. If $\lambda_{Ai} = \lambda_{Bi}$ for $i = 2, \dots, n$ all of the terms in (34) are zero so $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m = 0$.

As in Theorem 7, if $\lambda_{Ai} = \lambda_{Bi}$ for some $i \in \{2, \dots, n\}$, we cannot exclude the possibility that y_i happens to be the necessary nonzero value among $y_2, \dots, y_n$ while $y_j = 0$ for all $j \neq 1, i$, in which case all of the terms in (34) are zero so $\mathrm{d}r(\mathbf{M}(m)\mathbf{D})/\mathrm{d}m = 0$. Thus inequalities in 1 and 2 become non-strict if there is a single equality between non-Perron eigenvalues of $\mathbf{A}$ and $\mathbf{B}$. □

Remark 12. It is notable here that the relation on the Perron root, $r(\mathbf{AD}) > r(\mathbf{BD})$, is determined by the *non*-Perron eigenvalues of $\mathbf{A}$ dominating those of $\mathbf{B}$, for $\mathbf{A}$ and $\mathbf{B}$ with the same Perron root and nonscalar $\mathbf{D}$. Domination here means $\lambda_{Ai} > \lambda_{Bi}$ where i is the index on the non-Perron eigenvectors as ordered in $\mathbf{K}$. One may ask whether this relation holds if $\mathbf{K}$ is merely invertible rather than orthogonal as required for simultaneous symmetrizability of $\mathbf{A}$ and $\mathbf{B}$.

3 Applications

The inequalities examined here arise naturally in models of population dynamics. Karlin derived Theorem 2 [15, Theorem 5.1] in order to analyze the protection of genetic diversity in a subdivided population where $\mathbf{M}$ is the matrix of dispersal probabilities between patches. He wished to establish a partial ordering of stochastic matrices $\mathbf{M}$ with respect to their levels of 'mixing' over which $r(\mathbf{MD})$ decreases with increased mixing. Another partial ordering of matrices

examined by Karlin is of the form $\mathbf{M}(m) = (1-m)\mathbf{I} + m\mathbf{P}$ with mixing parameter m. Karlin's Theorem 5.2 [15] shows that $r(\mathbf{MD})$ decreases with increasing m, strictly when $\mathbf{P}$ is an irreducible stochastic matrix and $\mathbf{D}$ is a nonscalar positive diagonal matrix.

Variation in $(1-m)\mathbf{I} + m\mathbf{P}$ over m represents variation in the incidence of a single transforming processes (such as mutation, recombination, or dispersal) that scales all transitions between states equally. However, many natural systems have multiple transforming processes that act simultaneously, in which case the variation with respect to a single one of these processes generally takes the form $(1-m)\mathbf{Q} + m\mathbf{P}$ where $\mathbf{Q}$ is also a stochastic matrix. Karlin's Theorem 5.2 does not apply for general $\mathbf{Q} \neq \mathbf{I}$. The motivation to develop Theorem 17 [2, Theorem 2], below, was to extend Karlin's Theorem 5.2 to processes with multiple transforming events.

An open problem posed in [1] and [5] is the general characterization of the matrices $\mathbf{Q}$, $\mathbf{P}$, and $\mathbf{D}$ such that $r([(1-m)\mathbf{Q} + m\mathbf{P}]\mathbf{D})$ strictly decreases in m. Theorem 3 [3, Theorem 33] goes part way toward this characterization.

3.1 Temporal Properties

Theorem 3 was obtained to generalize a model by McNamara and Dall [18] of a population that disperses in a field of sites undergoing random change between two environments, where each environment produces its own rate of population growth. In the generalization [3], the environments are modeled as a reversible Markov chain with transition matrix $\mathbf{P}$, and $\mathbf{Q} = \lim_{t\to\infty} \mathbf{P}^t$. For $\mathbf{P}$ to have all negative non-Perron eigenvalues means that the environments change almost completely from one time increment to the next, while all positive eigenvalues correspond to more moderate change.

The correspondence originally discovered by McNamara and Dall [18] was between the sojourn times of the random environmental states, and whether natural selection was for or against dispersal. The generalization of their model shows that the direction of evolution of dispersal and the sojourn times of the environment are both determined by conditions C1 and C2 on the signs of the non-Perron eigenvalues of the environmental change matrix [3, Theorem 33]. More specifically, it is an inequality on the harmonic mean of the expected sojourn times of the Markov chain that is determined by conditions C1 and C2. The inequality derives from a remarkably little-known identity.

Lemma 13 (Harmonic Mean of Sojourn Times [3, Lemma 32])**.** *For a Markov chain with transition matrix $\mathbf{P}$, let $\tau_i(\mathbf{P})$ be the expected* sojourn time *in i (the mean duration of state i), and let $\{\lambda_i(\mathbf{P})\}$ be the eigenvalues of $\mathbf{P}$. Let E_A and E_H represent the unweighted arithmetic and harmonic means, respectively.*

These are related by the following identities:

$$E_H(\tau_i(\mathbf{P}))\left(1 - E_A(\lambda_i(\mathbf{P})\right) = 1, \tag{35}$$

or equivalently

$$E_A(\lambda_i(\mathbf{P})) + \frac{1}{E_H(\tau_i(\mathbf{P}))} = 1. \tag{36}$$

I should qualify "little known" — a version of (35) is well-known within the field of research on social mobility, but no reference to it outside this community appears evident. The identity arises in Shorrock's [22] social mobility index

$$\hat{M}(\mathbf{P}) = \frac{1}{n-1}\sum_{i=1}^{n}(1 - P_{ii}),$$

where P_{ij} is the probability of transition from social class j to class i. Shorrocks notes that $\hat{M}(\mathbf{P})$ is related to the expected sojourn times ('exit times') for each class i, $\tau_i = 1/(1 - P_{ii})$, through their harmonic mean,

$$E_H(\tau_i) := \frac{1}{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\frac{1}{\tau_i}} .$$

Evaluation gives

$$E_H\left(\frac{1}{1-P_{ii}}\right) = \frac{1}{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\frac{1}{1/(1-P_{ii})}} = \frac{n}{\displaystyle\sum_{i=1}^{n}(1-P_{ii})},$$

yielding

$$\hat{M}(\mathbf{P}) = \frac{1}{n-1}\sum_{i=1}^{n}(1 - P_{ii}) = \left(\frac{n}{n-1}\right)\frac{1}{E_H\left(\frac{1}{1-P_{ii}}\right)}.$$

Geweke et al. [12] define another social mobility index,

$$M_E(\mathbf{P}) = \frac{n - \sum_{i=1}^{n}|\lambda_i(\mathbf{P})|}{n-1}.$$

They note that when all the eigenvalues of $\mathbf{P}$ are real and nonnegative, then $\hat{M}(\mathbf{P}) = M_E(\mathbf{P})$, by the trace identity $\sum_{i=1}^{n} P_{ii} = \sum_{i=1}^{n}\lambda_i(\mathbf{P})$. Numerous papers cite this correspondence [19, 20]. However no expression of the identity in terms of the harmonic and arithmetic means, as in the forms (35) or (36), is evident.

Next, the eigenvalue conditions C1 and C2 are applied to the identity (35).

Theorem 14 (From [3, Theorem 33])**.** *Let $\mathbf{P}$ be the $n \times n$ transition matrix of an irreducible Markov chain whose eigenvalues are real. Let $\tau_i(\mathbf{P}) = 1/(1 - P_{ii})$ be the expected sojourn times in state i.*

C1. If all eigenvalues of $\mathbf{P}$ *are positive, then*

$$E_H(\tau_i(\mathbf{P})) > 1 + \frac{1}{n-1}. \tag{37}$$

C2. If all non-Perron eigenvalues of $\mathbf{P}$ *are negative, then*

$$1 \leq E_H(\tau_i(\mathbf{P})) < 1 + \frac{1}{n-1}. \tag{38}$$

C3. If all non-Perron eigenvalues of $\mathbf{P}$ *are zero, then*

$$1 \leq E_H(\tau_i(\mathbf{P})) = 1 + \frac{1}{n-1}. \tag{39}$$

C4. If all non-Perron eigenvalues of $\mathbf{P}$ *are the same sign or zero, and at least one is nonzero, then inequalities* (37) *and* (38) *are unchanged.*

Proof. The following equivalent inequalities are readily derived from one another:

$$E_H(\tau_i) = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\tau_i}} > 1 + \frac{1}{n-1} = \frac{n}{n-1} \tag{40}$$

$$\iff n-1 > \sum_{i=1}^{n}\frac{1}{\tau_i} = \sum_{i=1}^{n}\frac{1}{1/(1-P_{ii})} = n - \sum_{i=1}^{n} P_{ii}$$

$$\iff 1 < \sum_{i=1}^{n} P_{ii} = \sum_{i=1}^{n}\lambda_i(\mathbf{P}) = 1 + \sum_{i=2}^{n}\lambda_i(\mathbf{P})$$

$$\iff 0 < \sum_{i=2}^{n}\lambda_i(\mathbf{P}). \tag{41}$$

The analogous equivalence holds if the directions of the inequalities are reversed. If $\lambda_i(\mathbf{P}) > 0$ for all i then (41), (40), and (37) hold. Conversely, if $\lambda_i(\mathbf{P}) < 0$ for $i = 2, \ldots, n$ then $\sum_{i=2}^{n}\lambda_i(\mathbf{P}) < 0$, reversing the direction of the inequalities, and the right side of (38) holds. The left side of (38)clearly holds since $\tau_i(\mathbf{P}) = 1/(1-P_{ii}) \geq 1$ for each i. If $\lambda_i(\mathbf{P}) = 0$ for $i = 2, \ldots, n$ then clearly $E_H(\tau_i) = 1 + 1/(n-1)$. If $\lambda_i(\mathbf{P}) \geq 0$ for $i \in \{2, \ldots, n\}$, and $\lambda_i(\mathbf{P}) > 0$ for some $i \in \{2, \ldots, n\}$, then $\sum_{i=2}^{n}\lambda_i(\mathbf{P}) > 0$ so (41) continues to hold; analogously for the reverse inequality. ☐

We have seen that conditions C1 and C2 are sufficient to determine opposite directions of inequality in two very different expressions, one involving the temporal behavior of a Markov chain, $E_H(\tau_i(\mathbf{P})) > 1 + 1/(n-1)$, and the other involving the interaction of the chain with heterogeneous growth rates, $\mathrm{d}r(\mathbf{P}[(1-m)\mathbf{I} + m\mathbf{Q}]\mathbf{D})/\mathrm{d}m < 0$ and $r(\mathbf{PD}) > r(\mathbf{P}^2\mathbf{D})$ (under condition C1; the reverse under C2).

The inference in these results goes from the eigenvalue sign conditions, C1 and C2, to the inequalities. The converse, an implication from the inequality directions to the eigenvalue sign conditions, is found only in the case $n = 2$. It would be of empirical interest to know if there exist classes of stochastic matrices $\mathbf{P}$ for $n \geq 3$ in which the temporal behavior has direct implications upon the spectral radius, i.e. $E_H(\tau_i(\mathbf{P}))$ tells us about $\mathrm{d}r(\mathbf{P}[(1-m)\mathbf{I} + m\mathbf{Q}]\mathbf{D})/\mathrm{d}m$ and $r(\mathbf{P}^2\mathbf{D})/r(\mathbf{PD})$, or vice versa, without recourse to conditions C1 and C2.

An example of such a class for $n \geq 3$ is devised using rank-one matrices. Let $\mathcal{P}_n$ be the set of probability vectors of length n, so $\mathbf{e}^\top \mathbf{x} = 1, \mathbf{x} \geq \mathbf{0}$ for $\mathbf{x} \in \mathcal{P}_n$. Define the set of stochastic matrices

$$\mathcal{R}_n := \left\{ (1-\alpha)\mathbf{I} + \alpha \mathbf{v}\,\mathbf{e}^\top : \mathbf{v} \in \mathcal{P}_n, \mathbf{v} > \mathbf{0}, \alpha \in \big(0, \min_i \frac{1}{1-v_i}\big] \right\}.$$

Corollary 15. *Let $\mathbf{P} \in \mathcal{R}_n$, let $\mathbf{Q}$ be a symmetrizable irreducible stochastic matrix that commutes with $\mathbf{P}$, and let $\mathbf{D}$ be an $n \times n$ nonscalar positive diagonal matrix. Then*

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{P}[(1-m)\mathbf{I} + m\mathbf{Q}]\mathbf{D}) < 0 \textit{ if and only if } E_H(\tau_i(\mathbf{P})) > 1 + \frac{1}{n-1},$$
$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{P}[(1-m)\mathbf{I} + m\mathbf{Q}]\mathbf{D}) = 0 \textit{ if and only if } E_H(\tau_i(\mathbf{P})) = 1 + \frac{1}{n-1},$$

and

$$\frac{\mathrm{d}}{\mathrm{d}m} r(\mathbf{P}[(1-m)\mathbf{I} + m\mathbf{Q}]\mathbf{D}) > 0 \textit{ if and only if } 1 \leq E_H(\tau_i(\mathbf{P})) < 1 + \frac{1}{n-1}.$$

Corollary 16. *Let $\mathbf{P} \in \mathcal{R}_n$, and let $\mathbf{D}$ be an $n \times n$ nonscalar positive diagonal matrix. Then*

$$r(\mathbf{P}^2\mathbf{D}) < r(\mathbf{PD}) \textit{ if and only if } E_H(\tau_i(\mathbf{P})) > 1 + \frac{1}{n-1},$$
$$r(\mathbf{P}^2\mathbf{D}) = r(\mathbf{PD}) \textit{ if and only if } E_H(\tau_i(\mathbf{P})) = 1 + \frac{1}{n-1},$$

and

$$r(\mathbf{P}^2\mathbf{D}) > r(\mathbf{PD}) \textit{ if and only if } 1 \leq E_H(\tau_i(\mathbf{P})) < 1 + \frac{1}{n-1}.$$

Proof. The upper bound on α is the largest that assures $A_{ii} = 1 - \alpha + \alpha v_i \geq 0$ for each i. In addition it gives

$$1 - \alpha \in \big[1 - \min_i \frac{1}{1-v_i}, 1\big) = \big[-\min_i \frac{v_i}{1-v_i}, 1\big). \tag{42}$$

Any $\mathbf{P} \in \mathcal{R}_n$ is irreducible since by hypothesis $\mathbf{v} > \mathbf{0}$, $\alpha > 0$. To apply Theorem 7, we must verify that $\mathbf{P} \in \mathcal{R}_n$ is symmetrizable:

$$\mathbf{P} = (1-\alpha)\mathbf{I} + \alpha \mathbf{v}\,\mathbf{e}^\top = \mathbf{D}_{\mathbf{v}}^{1/2}[(1-\alpha)\mathbf{I} + \alpha(\mathbf{D}_{\mathbf{v}}^{1/2}\mathbf{e}\,\mathbf{e}^\top\mathbf{D}_{\mathbf{v}}^{1/2})]\mathbf{D}_{\mathbf{v}}^{-1/2}.$$

Let $\mathbf{z}_i$ be a right eigenvector of $\mathbf{P} \in \mathcal{R}_n$ associated with $\lambda_i(\mathbf{P})$. Then

$$\begin{aligned}\lambda_i \mathbf{z}_i = \mathbf{P}\mathbf{z}_i &= [(1-\alpha)\mathbf{I} + \alpha \mathbf{v}\,\mathbf{e}^\top]\mathbf{z}_i = (1-\alpha)\mathbf{z}_i + \alpha\mathbf{v}(\,\mathbf{e}^\top \mathbf{z}_i)\\ &\iff (\lambda_i - 1 + \alpha)\mathbf{z}_i = \alpha \mathbf{v}(\,\mathbf{e}^\top\mathbf{z}_i)\\ &\iff \mathbf{e}^\top\mathbf{z}_i = 0, \lambda_i = 1-\alpha \text{ or } \mathbf{z}_i = \frac{\alpha(\,\mathbf{e}^\top\mathbf{z}_i)}{\lambda_i+\alpha-1}\mathbf{v} = \mathbf{v}, \lambda_i = 1.\end{aligned}$$

So either (1) $\mathbf{z}_i$ is the right Perron vector of $\mathbf{P}$, or (2) $\lambda_i = 1-\alpha \in [-\min_j \frac{v_j}{1-v_j}, 1)$, by (42). Thus all of the non-Perron eigenvalues of $\mathbf{P}$ are equal, and may be either positive, zero, or negative in the range $[-\min_i \frac{v_i}{1-v_i}, 0)$, in which case exactly one of conditions C1, C3, or C2 is met, respectively, for Theorems 7 and 14 and Corollary 10, with the consequent implications. $\square$

3.2 Kronecker Products

A notable class of matrices that exhibit the commuting property required for Theorems 2, 3, 6, and 7 is the Kronecker product of powers of matrices. Define a set of square matrices

$$\mathcal{C} := \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_L\},$$

where each $\mathbf{A}_i$ is an $n_i \times n_i$ matrix. Define

$$\mathbf{M}(\mathbf{t}) := \bigotimes_{i=1}^{L} \mathbf{A}_i^{t_i} = \mathbf{A}_1^{t_1} \otimes \mathbf{A}_2^{t_2} \otimes \cdots \otimes \mathbf{A}_L^{t_L}, \tag{43}$$

where $\otimes$ is the Kronecker product (a.k.a. tensor product), $t_i \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}$, and $\mathbf{t} \in {\mathbb{N}_0}^L$. Now define the family of such products:

$$\mathcal{F}(\mathcal{C}) = \left\{ \bigotimes_{i=1}^{L} \mathbf{A}_i^{t_i} : t_i \in \{0, 1, 2, \ldots\} \right\}.$$

Clearly, any two members of $\mathcal{F}(\mathcal{C})$ commute, because for any $\mathbf{p}, \mathbf{q} \in {\mathbb{N}_0}^L$, then

$$\mathbf{M}(\mathbf{p})\,\mathbf{M}(\mathbf{q}) = \bigotimes_{i=1}^{L} \mathbf{A}_i^{p_i+q_i} = \mathbf{M}(\mathbf{q})\,\mathbf{M}(\mathbf{p}).$$

Products of the form $\bigotimes_{i=1}^{L} \mathbf{A}_i^{t_i}$ arise in multivariate Markov chains for which each variate X_i constitutes an independent Markov chain with transition matrix $\mathbf{A}_i$. The joint Markov process is exemplified be the transmission of information in a string of L symbols where transmission errors occur independently for each symbol. Such a process includes the genetic transmission of DNA or RNA sequences with independent mutations at each site. Under mitosis, the genome replicates approximately according to a transition matrix for a string of symbols with independent transmission errors at each site i:

$$\mathbf{M}_{\mathbf{m}} := \bigotimes_{i=1}^{L} [(1-m_i)\mathbf{I}_i + m_i \mathbf{P}_i] = \mathbf{A}[(1-m_k)\mathbf{I} + m_k\mathbf{B}],$$

where m_i is the probability of a transforming event at site i, and $\mathbf{P}_i$ is the transition matrix for site i given that a transforming event has occurred there. The form $\mathbf{A}[(1-m_k)\mathbf{I}+m_k\mathbf{B}]$ is provided to show the relationship to Theorem 7, where k may be any choice in $\{1,\ldots,L\}$, with

$$\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \cdots \otimes \mathbf{A}_{k-1} \otimes \ \mathbf{I}_k \ \otimes \mathbf{A}_{k+1} \otimes \cdots \otimes \mathbf{A}_L, \tag{44}$$

where $\mathbf{A}_i = (1-m_i)\mathbf{I}_i + m_i\mathbf{P}_i$, and

$$\mathbf{B} = \ \mathbf{I}_1 \ \otimes \mathbf{I}_2 \ \otimes \cdots \otimes \ \mathbf{I}_{k-1} \ \otimes \mathbf{P}_k \otimes \ \mathbf{I}_{k+1} \ \otimes \cdots \otimes \ \mathbf{I}_L. \tag{45}$$

However, both $\mathbf{A}$ and $\mathbf{B}$ in (44) and (45) are reducible due to the $\mathbf{I}$ terms, and this somewhat alters Theorem 7's condition on $\mathbf{D}$ for strict monotonicity of spectral radius. This is seen in Theorem 17.

Theorem 17. *[2, Theorem 2] Consider the stochastic matrix*

$$\mathbf{M}_{\mathbf{m}} := \bigotimes_{\xi=1}^{L} \left[(1-m_\xi)\mathbf{I}_\xi + m_\xi \mathbf{P}_\xi\right], \tag{46}$$

where each $\mathbf{P}_\xi$ is an $n_\xi \times n_\xi$ transition matrix for a reversible aperiodic Markov chain, $\mathbf{I}_\xi$ the $n_\xi \times n_\xi$ identity matrix, $L \geq 2$, and $\mathbf{m} \in (0, 1/2)^L$. Let $\mathbf{D}$ be a positive $N \times N$ diagonal matrix, where $N := \prod_{\xi=1}^{L} n_\xi$.

Then for every point $\mathbf{m} \in (0,1/2)^L$, the spectral radius of

$$\mathbf{M}_{\mathbf{m}}\mathbf{D} = \left\{ \bigotimes_{\xi=1}^{L} \left[(1-m_\xi)\mathbf{I}_\xi + m_\xi \mathbf{P}_\xi\right] \right\} \mathbf{D}$$

is non-increasing in each m_ξ.

If diagonal entries

$$D_{i_1 \cdots i_\xi \cdots i_L}, \ D_{i_1 \cdots i'_\xi \cdots i_L} \tag{47}$$

differ for at least one pair $i_\xi, i'_\xi \in \{1,\ldots,n_\xi\}$, for some $i_1 \in \{1,\ldots,n_1\}$, ..., $i_{\xi-1} \in \{1,\ldots,n_{\xi-1}\}$, $i_{\xi+1} \in \{1,\ldots,n_{\xi+1}\}$, ..., $i_L \in \{1,\ldots,n_L\}$, then

$$\frac{\partial r(\mathbf{M}_{\mathbf{m}}\mathbf{D})}{\partial \mu_\xi} < 0.$$

Remark 18. The condition on the diagonal entries (47) can be simply expressed in the cases $\xi = 1$ and $\xi = L$, respectively, as a requirement that $\mathbf{D} \neq c\,\mathbf{I}_1 \otimes \mathbf{D}'$ and $\mathbf{D} \neq \mathbf{D}' \otimes c\,\mathbf{I}_L$ for any $c \in \mathbb{R}$ and any $\mathbf{D}'$. For $\xi \in \{2,\ldots,L-1\}$ similar expressions could be given by employing permutations of the tensor indices.

Theorem 17 was obtained to characterize the effect of mutation rates on a clonal population, or on a gene that modifies mutation rates in a non-recombining

genome. This theorem shows that the asymptotic growth rate of an infinite population of types $\{(i_1, i_2, \ldots, i_L)\}$ is a strictly decreasing function of each mutation rate m_ξ when the growth rates D_i in (47) differ, and non-increasing otherwise. All the eigenvalues of $\mathbf{M}_\mathbf{m}$ are positive, as in condition C1 in Theorem 7, due to the assumption $m_\xi < 1/2$ for $\xi \in \{1, \ldots, L\}$.

The asymptotic growth rate of a quasispecies [11] at a mutation-selection balance is thus shown by Theorem 17 to be a decreasing function of the mutation rate for each base pair, a result not previously obtained with this level of generality in the multilocus mutation parameters, mutation matrices, and multilocus selection coefficients. As a practical matter, however, in genetics L may be very large, for example $L \approx 6 \times 10^9$ for the human genome. For such large L, populations cannot exhibit the Perron vector as a stationary distribution since the population size is infinitesimal compared to the genome space of $n = 4^L$. However, in large populations models that examine a small-L portion or approximation of the full genome, the Perron vector may become relevant as the stationary distribution under selection and mutation.

Proposition 19. *Theorem 17 extends to general symmetrizable irreducible nonnegative matrices*

$$\mathbf{M}_\mathbf{m} = \bigotimes_{\xi=1}^{L} \left[(1 - m_\xi) r(\mathbf{A}_\xi)\mathbf{I}_\xi + m_\xi \mathbf{A}_\xi\right],$$

where each $\mathbf{A}_\xi$ is a symmetrizable irreducible nonnegative $n_\xi \times n_\xi$ matrix.

Proof. For any given $\mathbf{A}$, let ${\mathbf{u}_A}^\top$ be its left Perron vector, and define

$$\mathbf{P} = \frac{1}{r(\mathbf{A})} \mathbf{D}_{\mathbf{u}_A} \mathbf{A}\, \mathbf{D}_{\mathbf{u}_A}^{-1}. \tag{48}$$

Then $\mathbf{P}$ is a symmetrizable irreducible stochastic matrix:

1. $\mathbf{P} \geq \mathbf{0}$ since $r(\mathbf{A}) > 0$, and $\mathbf{u}_A > \mathbf{0}$.
2. $\mathbf{P}$ is stochastic, since
$$\mathbf{e}^\top \left(\frac{1}{r(\mathbf{A})} \mathbf{D}_{\mathbf{u}_A} \mathbf{A} \mathbf{D}_{\mathbf{u}_A}^{-1} \right) = \frac{1}{r(\mathbf{A})} {\mathbf{u}_A}^\top \mathbf{A} \mathbf{D}_{\mathbf{u}_A}^{-1} = \frac{r(\mathbf{A})}{r(\mathbf{A})} {\mathbf{u}_A}^\top \mathbf{D}_{\mathbf{u}_A}^{-1} = \mathbf{e}^\top.$$
3. $\mathbf{P}$ is symmetrizable:
$$\mathbf{P} = \frac{1}{r(\mathbf{A})} \mathbf{D}_{\mathbf{u}_A} \mathbf{A} \mathbf{D}_{\mathbf{u}_A}^{-1} = \frac{1}{r(\mathbf{A})} \mathbf{D}_{\mathbf{u}_{A_\xi}} \mathbf{D}_\mathsf{L} \mathbf{S} \mathbf{D}_\mathsf{R} \mathbf{D}_{\mathbf{u}_A}^{-1} = \mathbf{D}'_\mathsf{L} \mathbf{S} \mathbf{D}'_\mathsf{R},$$
where
$$\mathbf{D}'_\mathsf{L} = \frac{1}{r(\mathbf{A})} \mathbf{D}_{\mathbf{u}_A} \mathbf{D}_\mathsf{L}, \qquad \mathbf{D}'_\mathsf{R} = \mathbf{D}_\mathsf{R} \mathbf{D}_{\mathbf{u}_A}^{-1}.$$

4. $\mathbf{P}$ is irreducible since $[\mathbf{A}^t]_{ij} > 0$ if and only if
$$[\mathbf{P}^t]_{ij} = r(\mathbf{A})^{-t}\ u_{Ai}\ u_{Aj}^{-1}\ [\mathbf{A}^t]_{ij} > 0.$$

The spectral radius expression with $\mathbf{A}_\xi$ terms is now shown to be equivalent to one with $\mathbf{P}_\xi$ terms:

$$\begin{aligned}
r(\mathbf{M}_\mathbf{m}\mathbf{D}) &= r\left(\bigotimes_{\xi=1}^{L}[(1-m_\xi)r(\mathbf{A}_\xi)\mathbf{I}_\xi + m_\xi\mathbf{A}_\xi]\,\mathbf{D}\right)\\
&= \prod_{\xi=1}^{L} r(\mathbf{A}_\xi)\ r\left(\bigotimes_{\xi=1}^{L}\left[(1-m_\xi)\mathbf{I}_\xi + m_\xi\frac{1}{r(\mathbf{A}_\xi)}\mathbf{A}_\xi\right]\mathbf{D}\right)\\
&= \prod_{\xi=1}^{L} r(\mathbf{A}_\xi)\ r\left(\bigotimes_{\xi=1}^{L}\left\{\mathbf{D}_{\mathbf{u}_{A\xi}}[(1-m_\xi)\mathbf{I}_\xi + m_\xi\frac{1}{r(\mathbf{A}_\xi)}\mathbf{A}_\xi]\mathbf{D}_{\mathbf{u}_{A\xi}}^{-1}\right\}\mathbf{D}\right)\\
&= \prod_{\xi=1}^{L} r(\mathbf{A}_\xi)\ r\left(\bigotimes_{\xi=1}^{L}\left[(1-m_\xi)\mathbf{I}_\xi + m_\xi\frac{1}{r(\mathbf{A}_\xi)}\mathbf{D}_{\mathbf{u}_{A\xi}}\mathbf{A}_\xi\mathbf{D}_{\mathbf{u}_{A\xi}}^{-1}\right]\mathbf{D}\right)\\
&= \prod_{\xi=1}^{L} r(\mathbf{A}_\xi)\ r\left(\bigotimes_{\xi=1}^{L}[(1-m_\xi)\mathbf{I}_\xi + m_\xi\mathbf{P}_\xi]\,\mathbf{D}\right)\\
&= \prod_{\xi=1}^{L} r(\mathbf{A}_\xi)\ r(\mathbf{M}'_\mathbf{m}\mathbf{D}),
\end{aligned}$$

where $\mathbf{M}'_\mathbf{m} := \bigotimes_{\xi=1}^{L}[(1-m_\xi)\mathbf{I}_\xi + m_\xi\mathbf{P}_\xi]$. Therefore

$$\frac{\partial}{\partial m_i} r(\mathbf{M}_\mathbf{m}\mathbf{D}) = \prod_{\xi=1}^{L} r(\mathbf{A}_\xi)\ \frac{\partial}{\partial m_i} r(\mathbf{M}'_\mathbf{m}\mathbf{D}). \tag{49}$$

Theorem 17, being applicable to the right hand side of (49), is thus extended to the left hand side involving general symmetric irreducible nonnegative matrices. □

Remark 20. The identification $\mathbf{P} = \frac{1}{r(\mathbf{A})}\mathbf{D}_{\mathbf{u}_A}\mathbf{A}\mathbf{D}_{\mathbf{u}_A}^{-1}$ (48) used in Proposition 19 also provides another route to extend Theorem 3 to Theorem 7, sidestepping Lemmas 4 and 5, Theorem 6, and the proof of Theorem 7. But analogous steps are the basis for the proof of Theorem 3 in [3], so they are provided here to be self-contained.

3.3 Other Applications

Condition C2 is met by nonnegative *conditionally negative definite* matrices [6, Chapter 4], [7]. Symmetric conditionally negative definite matrices arise in the analysis of the one-locus, multiple-allele viability selection model. If the

matrix of fitness coefficients $\mathbf{W}$ allows the existence of a polymorphism with all alleles present, then the polymorphism is globally stable if $\mathbf{W}$ is conditionally negative definite [17] (Kingman's exact condition being that they need only be conditionally negative semidefinite).

4 Open Problems

The conditions in Theorem 7 that all the eigenvalues of $\mathbf{A}$ be positive (C1), or that all the non-Perron eigenvalues be negative (C2), are clearly very strong, and leave us with no results for intermediate conditions. One can no doubt obtain intermediate results by placing additional conditions on the matrices $\mathbf{E}$, $\mathbf{K}$, and $\mathbf{D}$, but this remains an unexplored area. However, it is seen that for $n = 2$ the conditions C1 and C2 cannot be weakened.

The condition of symmetrizability also imposes a large constraint on the generality of the results here. The inequalities described undoubtedly have extensions to general nonnegative matrices. For the more general case, however, we lose use of the Rayleigh-Ritz variational formula, and the spectral radius no longer has the sum-of-squares representation (13), which is our principal tool. The general case remains an open question.

Acknowledgements

I thank Joel E. Cohen for bringing to my attention the question of $r(\mathbf{A})r(\mathbf{AD})$ vs. $r(\mathbf{A}^2\mathbf{D})$, and I thank Shmuel Friedland for referring Joel Cohen to me.

References

[1] L. Altenberg. Open problems in the spectral analysis of evolutionary dynamics. In A. Menon, editor, *Frontiers of Evolutionary Computation*, volume 11 of *Genetic Algorithms And Evolutionary Computation Series*, pages 73–102. Kluwer Academic Publishers, Boston, MA, 2004.

[2] L. Altenberg. An evolutionary reduction principle for mutation rates at multiple loci. *Bulletin of Mathematical Biology*, 73:1227–1270, 2011.

[3] L. Altenberg. The evolution of dispersal in random environments and the principle of partial control. *Ecological Monographs*, 82(3):297–333, 2012.

[4] L. Altenberg. The principle of partial control in reaction-diffusion models for the evolution of dispersal. *Under review*, 2012.

[5] L. Altenberg. Resolvent positive linear operators exhibit the reduction phenomenon. *Proceedings of the National Academy of Sciences U.S.A.*, 109(10):3705–3710, 2012.

[6] R. B. Bapat and T. E. S. Raghavan. *Nonnegative Matrices and Applications*. Cambridge University Press, Cambridge, UK, 1997.

[7] R. Bhatia. *Positive Definite Matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007.

[8] H. Caswell. *Matrix Population Models*. Sinauer Associates, Sunderland, MA, 2nd edition, 2000.

[9] J. E. Cohen. Cauchy inequalities for the spectral radius of products of diagonal and nonnegative matrices. *Proceedings of the American Mathematical Society*, page In press, 2012.

[10] J. E. Cohen, S. Friedland, T. Kato, and F. P. Kelly. Eigenvalue inequalities for products of matrix exponentials. *Linear Algebra and Its Applications*, 45:55–95, 1982.

[11] M. Eigen and P. Schuster. The hypercycle: A principle of natural self-organization. *Naturwissenschaften*, 64:541–565, 1977.

[12] J. Geweke, R. Marshall, and G. Zarkin. Mobility indices in continuous time Markov chains. *Econometrica*, 54(6):1407–1423, 1986.

[13] S. H. Gould. *Variational Methods for Eigenvalue Problems: An Introduction to the Weinstein Method of Intermediate Problems*, volume 10. University of Toronto Press; London: Oxford University Press, 1966.

[14] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[15] S. Karlin. Classifications of selection–migration structures and conditions for a protected polymorphism. In M. K. Hecht, B. Wallace, and G. T. Prance, editors, *Evolutionary Biology*, volume 14, pages 61–204. Plenum Publishing Corporation, New York, 1982.

[16] J. Keilson. *Markov Chain Models: Rarity and Exponentiality*. Springer-Verlag, New York, 1979.

[17] J. F. C. Kingman. A mathematical problem in population genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 57:574–582, 1961.

[18] J. M. McNamara and S. R. Dall. The evolution of unconditional strategies via the 'multiplier effect'. *Ecology Letters*, 14(3):237–243, Mar. 2011.

[19] D. Quah. Aggregate and regional disaggregate fluctuations. *Empirical Economics*, 21(1):137–159, 1996.

[20] S. Redding. Specialization dynamics. *Journal of International Economics*, 58(2):299–334, 2002.

[21] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, revised edition, 2006.

[22] A. F. Shorrocks. The measurement of mobility. *Econometrica*, 46(5):1013–1024, 1978.